Computer Vision and Image Understanding
journal homepage: www.elsevier.com

# Curvature-based spectral signatures for non-rigid shape retrieval

Frederico A. Limberger[a,**], Richard C. Wilson[a]

[a] University of York, Department of Computer Science, York, YO10 5GH, United Kingdom

ABSTRACT

The geometric properties of descriptors derived from the diffusion geometry family have many valuable properties for shape analysis. These descriptors, also known as diffusion distances, use the eigenvalues and eigenfunctions of the Laplace-Beltrami operator to construct invariant metrics about the shape. Although they are invariant to many transformations, non-rigid deformations still modify the shape spectrum. In this paper, we propose a shape descriptor framework based on a Lagrangian formulation of dynamics on the surface of the object. We show how our framework can be applied to non-rigid shape retrieval, once it benefits from the analysis and the automatic identification of shape joints, using a curvature-based scheme to identify these regions. We also propose modifications to the Improved Wave Kernel Signature in order to keep descriptors more stable against non-rigid deformations. We compare our spectral components with the classic ones and our spectral framework with state-of-the-art non-rigid signatures on traditional benchmarks, showing that our shape spectra is more stable and discriminative and clearly outperforms other descriptors in the SHREC'10, SHREC'11 and SHREC'17 benchmarks.



## 1. Introduction

Designing feature descriptors is crucial in shape analysis. However, it is not a simple task to describe the important characteristics of the shape and still remain invariant to the complex transformations that the shape may undergo. Spectral descriptors have gained increased attention for their advantageous properties, mainly because they are intrinsically invariant to common shape deformations. For example, they are invariant to Euclidean transformations, which was the main focus of early research in this field (Belongie et al. (2000); Johnson (1997)), and relatively stable against non-rigid articulations, where significant attention has been given in the past decade. However, in most recent non-rigid benchmarks (Lian et al. (2010, 2011); Pickup et al. (2014); Lian et al. (2015)), spectral descriptors have not performed at their full potential. Shape articulations have proved to be hard to code into a descriptor since it is difficult to distinguish an articulated model from a model of a similar class.

In this paper, we present new methods for composing non-rigid shape signatures, which are more stable to non-rigid transformations, by computing enhanced spectral signatures from 3D meshes and by modifying how curvature is aggregated in the Improved Wave Kernel Signature (Limberger and Wilson (2015)). Our **main contributions** include:

- Kinetic Laplace-Beltrami Operator (KLBO): Kinetic spectral components computed from 3D triangle meshes. We compute spectral signatures by weighting the LBO by a curvature-based kinetic term. This weight removes the influence of shape's articulations on shape descriptors. The weighting is small in areas on the shape where articulations are likely to occur and also stable to rigid and non-rigid motions. We show how to modify the kinetic density in the eigensystem to generate a consistent spectrum to deformable objects. How we construct this new eigensystem is primordial to correctly take into account the modifications of the kinetic energy in the descriptor. More details are in Section 3.

- Improved Wave Kernel Signature (IWKS): We detail how to compute a better signature for 3D non-rigid shape retrieval based on a different energy scaling (Limberger and Wilson (2015)). Also, we show how to properly integrate extrinsic information to the signature to make it more

**Corresponding author: Tel.: +55 55999408899; fax: +44 01904 325599; *e-mail:* fredericoal@gmail.com (Frederico A. Limberger)

discriminative over the encoding process and more stable over shape scaling. More details are in Section 4.

- The integration of KLBO signatures with Fisher Vector and Super Vector encoding schemes, which use Gaussian Mixture Models (GMM) to create a feature dictionary. We show that IWKS signatures can be used with these state-of-the-art encoding schemes since IWKS descriptors can be precisely approximated by a probabilistic distribution function. We also compare the convergence against HKS, SIHKS and WKS showing that the IWKS presents similar approximation errors. More details are in Section 4.1.

The remainder of this paper is organized as follows. Section 2 describes previous related works in respect to non-rigid 3D shape descriptors. In Section 3, we introduce the enhanced spectral components based on classical field theory to reduce the influence of shape motions on descriptors. In Section 4, we detail how to properly integrate extrinsic information to the IWKS. Section 4.1 shows how we compute and apply state-of-the-art encoding schemes to KLBO descriptors. Finally, Section 5 shows experiments on the differences of the LBO to the KLBO and presents evaluation performances on a number of shape databases, comparing the KLBO signatures against state-of-the-art techniques.

## 2. Related Work

In this section, we review works related to different non-rigid shape descriptors proposed in the literature. We divided the shape descriptors into three categories: spectral-based, geometry-based, and learning-based descriptors.

### 2.1. Spectral-based descriptors

Spectral-based descriptors are based on solutions that rely on the analysis of the eigensystem (eigenvalues and eigenfunctions) of the Laplace-Beltrami operator. These spectral components have many interesting properties which can be combined to compose an elegant solution to the non-rigid shape retrieval problem. This technique was first applied to represent shapes by Reuter et al. (2005). They have used the eigenvalues of the LBO as shape fingerprints for shape identification and comparison. Right after, Rustamov (2007) created a global shape descriptor using the eigenvalues and eigenfunctions of the LBO to describe a object.

Following, Sun et al. (2009) proposed the Heat Kernel Signature (HKS) which is based on the diffusion of heat over the surface of the model, governed by the heat equation

$$\Delta_{\mathcal{M}} u(x,t) = \frac{\partial u}{\partial t}(x,t), \quad (1)$$

where $u$ is a function in respect to space and time that requires to satisfy the Dirichlet boundary condition $u(x,t) = 0$ for all $x \in \partial\mathcal{M}$ during all $t$ and $\mathcal{M}$ is a Riemannian manifold. Given a starting heat distribution at time $t$ the purpose is to measure how the heat is diffused across the shape to compute a heat-based descriptor. Later, Bronstein and Kokkinos (2010) created a framework to transform the HKS in a scale-invariant descriptor (SIHKS) (i.e. signature does not depend on the size of the shape). Then, Aubry et al. (2011) proposed the Wave Kernel Signature (WKS) based on the Schrödinger equation

$$i\Delta_{\mathcal{M}} \psi(x,t) = \frac{\partial \psi}{\partial t}(x,t), \quad (2)$$

which is very similar to the heat equation but it has different induced dynamics. Instead of using different time intervals, they compute the descriptor at different energy scales. By analyzing the eigenvalue distributions of same-class shapes, Limberger and Wilson (2015) proposed the Improved Wave Kernel signature which has a more informative scaling (power scaling) to the eigenvalues of the WKS. They also propose to use a curvature aggregation based on principal curvatures to improve histogram discrimination.

Recently, Ye and Yu (2015) proposed a framework specifically for encoding non-rigid geometries by using a context-aware integral kernel operator on a manifold, taking advantage of functional operators. Li and BenHamza (2013) introduced a spectral graph wavelet framework to retrieve shapes in non-rigid databases, using a multiresolution descriptor which can capture the global and local geometry of 3D shapes. Masoumi et al. (2016) improved the work of Li and BenHamza (2013) by incorporating the vertex area in the computation of the descriptor. Li et al. (2016) computed a descriptor for non-rigid shape retrieval based on the HKS which is only computed on assigned key-points to reduce computational complexity and increase descriptiveness. Mohamed and Ben 4 (2016) proposed a descriptor based on the spectral shape skeleton computed from the second eigenfunction of the LBO, and used a graph matching framework to compare skeletons.

### 2.2. Geometry-based descriptors

Geometry-based descriptors use statistics computed on primitive geometric attributes, for instance, distance between any two points and/or shape histograms, to characterize 3D models. Many techniques have been proposed to handle non-rigid shape retrieval.

One way of addressing this problem is by first applying multidimensional scaling (MDS) to transform models to their canonical form thus removing the influence of motions, then computing a shape description. Elad and Kimmel (2003) computed bending invariant signatures by applying MDS on the intrinsic geodesic distances between surface points, computed from the fast marching on triangulated domains algorithm. Lian et al. (2013) applied Clock Matching to depth-buffer images, captured around the 3D objects in their canonical forms (CM-BOF). To compute distances between models a multi-view shape matching is applied. Li et al. (2014a) proposed a hybrid descriptor (MDS-ZFDR) by combining MDS with distance-based and curvature-based features. Pickup et al. (2015) changed the way Lian et al. (2013) computed canonical forms by using Euclidean distances. Their algorithm has a similar accuracy but it has lower computation times. Pickup et al. (2016) computed a canonical form by unbending the skeleton of the mesh to perform non-rigid shape retrieval.

The second way of computing geometric descriptors is by extracting features from distinct views of 3D objects. In this sense, Furuya and Ohbuchi (2014) fused SIFTs, computed from views of the object, using an anchor manifold graph to create a more powerful descriptor. Later, Furuya and Ohbuchi (2015) developed a feature aggregation algorithm named *Diffusion-On-Manifold* to encode local features into a global descriptor. They also proposed a new local feature called *Position and Orientation Distribution*, that describes the oriented-points distribution using a *Sphere-Of-Interest*.

### *2.3. Learning-based descriptors*

Recently, machine learning methods, mainly convolution neural networks (CNN), have gained attention of researchers. In the shape retrieval field, there have been recent works concerned with applying supervised learning methods to geometric data. These methods require some prior knowledge such as training data so they can learn class attributes.

Litman et al. (2014) used a supervised construction of the dictionary in BoF learned via bi-level optimization. Later, Litman and Bronstein (2014) defined a framework to learn an optimized descriptor by taking into account the statistics of shapes. Boscaini et al. (2015) proposed a localized spectral CNN using the windowed Fourier transform to represent local shape structures and created class-specific descriptors for deformable shapes.

## 3. Kinetic Laplace-Beltrami operator

The family of spectral methods, exemplified by Sun et al. (2009) and Aubry et al. (2011) are very attractive for 3D shape representation because they are isometrically invariant, easy to make scale invariant, partly resistant to shape deformations, and easy to calculate even for large meshes. They are also resistant to some types of noise, which appears in the high-frequency part of the shape spectrum and can be downweighted. The essence of these methods is to define a dynamic equation on the surface of the shape (for example, the heat equation or the wave equation) and use the solution to extract information about the shape.

The Kinetic Laplace-Beltrami operator (KLBO) is an operator designed to reduce the influence of joint motions on shape descriptors. It does that by modifying the kinetic energy on the surface of the object, making energy more difficult to move in articulated regions. Before detailing the KLBO, we briefly review the classical Laplace-Beltrami operator and discuss some related works.

The Laplace operator generalized to operate on functions defined on a Riemannian manifold $\mathcal{M}$ (2D in our case) is known as the Laplace-Beltrami operator $\Delta_{\mathcal{M}}$. It is a linear operator defined as the divergence of the gradient taking functions into functions

$$\Delta_{\mathcal{M}} f = -\nabla \cdot \nabla_{\mathcal{M}} f \tag{3}$$

given that $f$ is a twice-differentiable real-valued function. The negative sign is simply to respect the standard convention for graph Laplacians.

Several methods were proposed to deal with the problem of creating discrete Laplacians for meshes: Taubin (1995); Desbrun et al. (1999); Meyer et al. (2003); Reuter et al. (2006); Xu (2006); Levy (2006); Belkin et al. (2008). Meyer et al. (2003) proposed the cotangent weight scheme which we use in our construction. Although Xu (2004) showed that the cotangent weight scheme does not converge in general, Belkin et al. (2008) proposed a new discretization method that converges even when meshes present imperfections. In this work, we did not have any convergence problems using Meyer et al. (2003), however, other discretization methods can also be applied to our method.

Take $f : \mathcal{V} \to \mathbb{R}$ as a $n$-dimensional vector where the $i$th component $f(i)$ is the function value at the vertex $i$ in $\mathcal{V}$. Using Meyer et al.'s discretization, the discrete Laplacian ($L$) is written as

$$\Delta_M = L = A^{-1} W \tag{4}$$

where $A$ is a positive definite diagonal matrix and the elements of W are given by

$$W(i,j) = \begin{cases} \frac{(\cot\alpha_{ij} + \cot\beta_{ij})}{2} & \text{if } (i,j) \in E, \\ -\sum\limits_{k \neq i} W(i,k) & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

where $\alpha_{ij}$ and $\beta_{ij}$ are internal angles ($\angle(\mathbf{v}_i\mathbf{v}_a\mathbf{v}_j)$ and $\angle(\mathbf{v}_i\mathbf{v}_b\mathbf{v}_j)$) of two adjacent triangles with center vertex $\mathbf{v}_i$ and $E$ is the edge set. The diagonal elements $A_{ii}$ are the Voronoi areas associated to the vertex $\mathbf{v}_i$. Figure 1 shows a diagram of this parameterization.

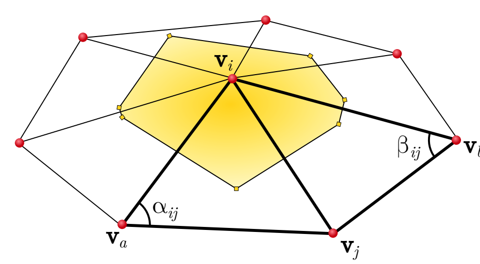


**Fig. 1. Illustration of the angles $\alpha_{ij}$ and $\beta_{ij}$, and the Voronoi area (yellow polygon) associated to the vertex $\mathbf{v}_i$ of the cotangent weight scheme of Meyer et al. (2003).**

Recently, Andreux et al. (2014) proposed an anisotropic LBO which benefits from a more semantically meaningful diffusion process, being able to favor directions of low or high curvature. Later, Boscaini et al. (2016) proposed shape descriptors constructed from anisotropic oriented diffusion kernels which take advantage of deep learning techniques.

Choukroun et al. (2016) proposed the use of the Hamiltonian $H = -\Delta + V$ and the corresponding Schrödinger equation $i\hbar\dot{\Psi} = H\Psi$ for shape analysis. They analyze the effect of the potential $V$ on the Hamiltonian eigenfunctions and show how $V$ can be optimized for particular representational problems. Our starting point is the Lagrangian and we use the Euler-Lagrange equations to obtain the dynamics. These are dual approaches; in the same physical situation they will lead to the same dynamics. The discretization proposed in Choukroun et al. (2016) is different to ours, as we add a kinetic energy related term to the Lagrangian rather than a potential.

The paper suggests that by choosing the right optimization of the potential it is possible to deal with different shape analysis tasks. They show how to deal with the problem of meshing compression by choosing a general optimization method for solving variational problems. Hamiltonian and Lagrangian mechanics are related formulations predicting the dynamics of a system. The Hamiltonian approach defines the time evolution of a system via a set of differential equations, whereas the Lagrangian framework proceeds from the principle of least action applied to the Lagrangian.

Similarly to Choukroun et al. (2016), Melzi et al. (2017) also have made a modification to the Laplacian via a potential function. They have created a new operator for computing localized manifold harmonics on deformable objects using the Laplacian eigenfunctions framework. Differently from ours, they focus on the local spectral shape analysis, which means constructing localized orthogonal bases that removes the global nature of these bases. On the other hand, we focus on constructing global shape representation that are invariant to local properties.

In this paper, we define the Lagrangian of the dynamics on the surface of objects using *classical field theory*. To be clear, this is different from a weighted manifold decomposition Grigor'yan (2006); Andreux et al. (2014). Different from Andreux et al. (2014), we do not simply favor directions of high or low curvature to create an anisotropic diffusion. Instead, we discount the kinetic term of joint regions, so that these do not influence the descriptors. Thus, we weight the physical field using a smooth positive kinetic density. Both methods share some functional similarities, however, our formulation for the problem and its outcome are completely different. We begin by defining the Lagrangian density of the system

$$\mathcal{L}(\phi, \triangledown\phi, \dot{\phi}, \mathbf{x}, t) = T - V \tag{6}$$

where $T$ is the kinetic energy (K.E.) and $V$ is the potential energy, and $\phi$ represents a field defined over the space (i.e. over the surface of the object). The *action* of the system is given by the integral of the Lagrangian density:

$$S(\mathcal{L}) = \int \mathcal{L} d\mathbf{x} dt \tag{7}$$

The dynamics of the system can be recovered from Hamilton's principle, which states that the action should be stationary for the true dynamic evolution of the system. This leads to the Euler-Lagrange equation for the dynamics:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \triangledown \cdot \frac{\partial \mathcal{L}}{\partial \triangledown \phi} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right) = 0 \tag{8}$$

By defining an appropriate Lagrangian and solving the resulting Euler-Lagrange equations, we can find a shape signature that weights kinetic energy differently across the field. The kinetic energy is generated by different forms of motions. The movement of joint regions can be physically interpreted as translational (when one part is moved from one place to another), rotational (when the joint is rotated) and/or vibrational (when part of the shape is also deformed by the motion), depending on the type of articulation and deformation. Thus, our motivation is to weight the kinetic energy over the shape surface to remove joint-articulation's effect on shape signatures.

In the general scaled Lagrangian:

$$\mathcal{L} = \frac{1}{2}\dot{\phi}\dot{\phi}^* + \frac{1}{2}(\phi\dot{\phi}^* - \phi^*\dot{\phi}) - \nabla\phi \cdot \nabla\phi^* - \phi^* V \phi \tag{9}$$

$\phi$ is a (possibly) complex field, so there are in fact two "fields" corresponding to the real part and the imaginary part. In practice it is easier to consider the field $\phi$ and its complex-conjugate $\phi^*$ which are linear combinations of the real and imaginary parts and so do not affect the calculations. The first two terms are kinetic energy terms, the first corresponding to a standard K.E. proportional to the square of the velocity. The second is a K.E. term where the energy increases with the size of the field. The second two terms are field potentials, the first related to the gradient and the second to some external potential field $V$.

From this point, we can define the shape descriptor by choosing appropriate terms from (9). To define the heat equation we choose second and third terms

$$\mathcal{L} = \frac{1}{2}(\phi\dot{\phi}^* - \phi^*\dot{\phi}) - \nabla\phi \cdot \nabla\phi^* \tag{10}$$

which applied to (8) gives dynamics

$$\nabla^2\phi = \dot{\phi}. \tag{11}$$

Ultimately this leads to the definition of the family of heat kernel signatures. In the same way, to define the wave equation we choose first and third terms

$$\mathcal{L} = \frac{1}{2}\dot{\phi}\dot{\phi}^* - \nabla\phi \cdot \nabla\phi^* \tag{12}$$

which gives dynamics

$$\nabla^2\phi = \ddot{\phi}. \tag{13}$$

which ultimately leads to the wave kernel signature of Aubry et al. (2011). The final term can be used to introduce a potential energy weighting term $V(\mathbf{x})$ which varies across the surface but we do not use that term here. To reduce the effect of object articulations, we introduce a kinetic energy weighting term $C(\mathbf{x})$ (first to the wave equation) into the Lagrangian

$$\mathcal{L} = \frac{k}{2}C(\mathbf{x})\dot{\phi}\dot{\phi}^* - \nabla\phi \cdot \nabla\phi^* \tag{14}$$

By applying this time (14) to (8) we get

$$\triangledown^2\phi = kC(\mathbf{x})\ddot{\phi} \tag{15}$$

where $k$ is a normalization term and $C(\mathbf{x})$ is a spatially varying weighting function which is small in areas on the shape where articulations are likely to occur and also stable to non-rigid motions. Standard separation of variables and discretization gives

$$L\phi = \lambda K\phi \tag{16}$$

where $K$ is a diagonal matrix where diagonal elements $K_{ii} = C(\mathbf{x})$ such that $i$ represents the vertex at position $\mathbf{x}$. Putting these elements together, following a standard discretization procedure from (4), the eigenvectors associated with the signature are solutions of the generalized eigenproblem

$$W\phi = \lambda A K \phi. \tag{17}$$

The kinetic term can be modified in the heat Lagrangian in the same way:

$$\mathcal{L} = \frac{k}{2} C(\mathbf{x})(\phi\dot{\phi}^* - \phi^*\dot{\phi}) - \nabla\phi \cdot \nabla\phi^* \tag{18}$$

and this leads to exactly the same spatial eigenvectors of (17), although the solution is different due to difference in the time derivatives. To derive a descriptor which is less variant to non-rigid motions we merely need to choose an appropriate function $C(\mathbf{x})$ which is smaller in articulated points than in rigid areas of the shape.

The weighting function $C(\mathbf{x})$ needs to reduce the effect of areas which are most different when comparing non-rigid shapes. When a human model moves its arm, what happens is that the arm joint region changes its curvature along with its local volume. At one side it becomes smaller (more negative) and at the other side it becomes bigger (more positive). After analyzing the structure of shapes we found a relation between the positive volume of the local surface patches and their joints. By using the positive volume, joint regions are consistently less weighted than other regions. We can compute the volume of a surface patch, similarly as done by Pottmann et al. (2009), by integrating a quadric patch, which is a representation of this surface in the local coordinate system

$$\iint_{x^2+y^2<R^2} (k_1 x^2 + k_2 y^2 - z)\,dxdy \tag{19}$$

Here $k_1$ and $k_2$ are the principal curvatures of the surface patch since they are the eigenvalues of the symmetric matrix in the second fundamental form

$$II = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} \tag{20}$$

Figure 3 shows a diagram of the volume of a surface patch. After integrations of (19), the volume inside the circle with radius $R$ centered at the respective vertex $\mathbf{v}_i$ is defined as

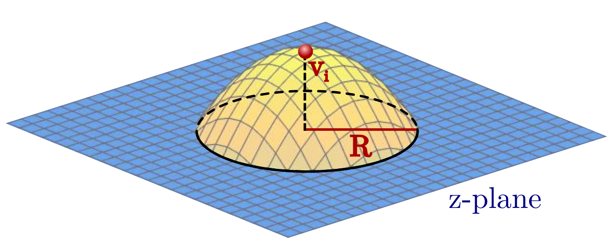


**Fig. 3. Diagram of the volume of a surface patch centered at vertex $\mathbf{v}_i$**

$$\frac{\pi R^4}{2}\frac{k_1 + k_2}{2} - \pi z R^2 \tag{21}$$

where $(k_1+k_2)/2$ is the **mean curvature** $(H)$, $\pi R^4/2$ is a scaling factor and $z$ dictates the position of the reference plane defined by the surface patch's normal. By removing the scaling factor and taking $z = 0$ we find that the volume is **proportional** to the mean curvature thus we define the weighting function $C(\mathbf{x})$ as

$$C(\mathbf{x}) = \max(\epsilon, H(\mathbf{x})), \tag{22}$$

where $H(\mathbf{x})$ is the mean curvature of the surface patch at position $\mathbf{x}$ and $\epsilon$ a very small number (e.g. $10^{-8}$). The scaling factor $\pi R^4/2$ of Eq. (21) is not significant because a scaling normalization is performed after this stage. Thus, to facilitate the volume computation, we directly extract mean curvatures from shapes by computing a curvature tensor at each vertex, according to Rusinkiewicz (2004).

With these ingredients, Eq. (16) can be solved as a generalized eigenvalue problem. The resulting eigensystem is then used to construct a shape signature following the appropriate method for the particular Lagrangian, i.e. using (14) for either the WKS or IWKS or (18) for either the HKS or SIHKS.

We call these methods, using modified kinetic energy terms in the Lagrangian, *Kinetic Laplace-Beltrami operator* or KLBO methods. Fig. 2 summarizes the main steps of the KLBO pipeline for computing spectral signatures.

## 4. Improved Wave Kernel Signature

The Improved Wave Kernel signature (IWKS), which was introduced by Limberger and Wilson (2015), presents a different energy scaling from the WKS of Aubry et al. (2011). This scaling is the result of an investigation of how the eigenvalues of the shapes modify after these being deformed by non-rigid motions. Further, the IWKS also incorporates extrinsic information into the descriptor to make the encoding more discriminative to other classes of objects. The WKS is given by

$$\mathrm{WKS}(x, e) = C_e \sum_{k=1}^{\infty} \phi_k(x)^2 f_E(\Lambda_k)^2 \tag{23}$$

$$f_E(\Lambda_k)^2 = e^{\frac{-(e-\log(\Lambda_k))^2}{2\sigma^2}} \qquad e \in [\log(\lambda_2), \log(\lambda_{\max})] \tag{24}$$

while the IWKS from Limberger and Wilson (2015) is given by

$$\mathrm{IWKS}(x, e) = C_e \sum_{k=1}^{\infty} \phi_k(x)^2 f_C(\Lambda_k)^2 + c_x\alpha \tag{25}$$

where $c_x$ is the maximum principal curvature, $\alpha$ is a weight that normalizes $c_x$ accordingly to the signature values and

$$f_C(\Lambda_k)^2 = e^{\frac{-(e-\sqrt[3]{\Lambda_k})^2}{2\sigma^2}} \qquad e \in [\sqrt[3]{\lambda_f}, \sqrt[3]{\lambda_{\max}}] \tag{26}$$

where $\lambda_f$ corresponds to the first non-zero eigenvalue.

The problem of using a constant $\alpha$ to balance the curvature term with the signatures values is that the curvatures will be different locally if the shape appears in a different scale or if the shape is deformed by its joints. Thus, we propose a new weighting term that takes these factors into consideration and weights the curvature terms in a way that it keeps the average curvature constant.

$$S(x) = \frac{\beta \cdot c_x}{\mathrm{mean}(c)} \tag{27}$$

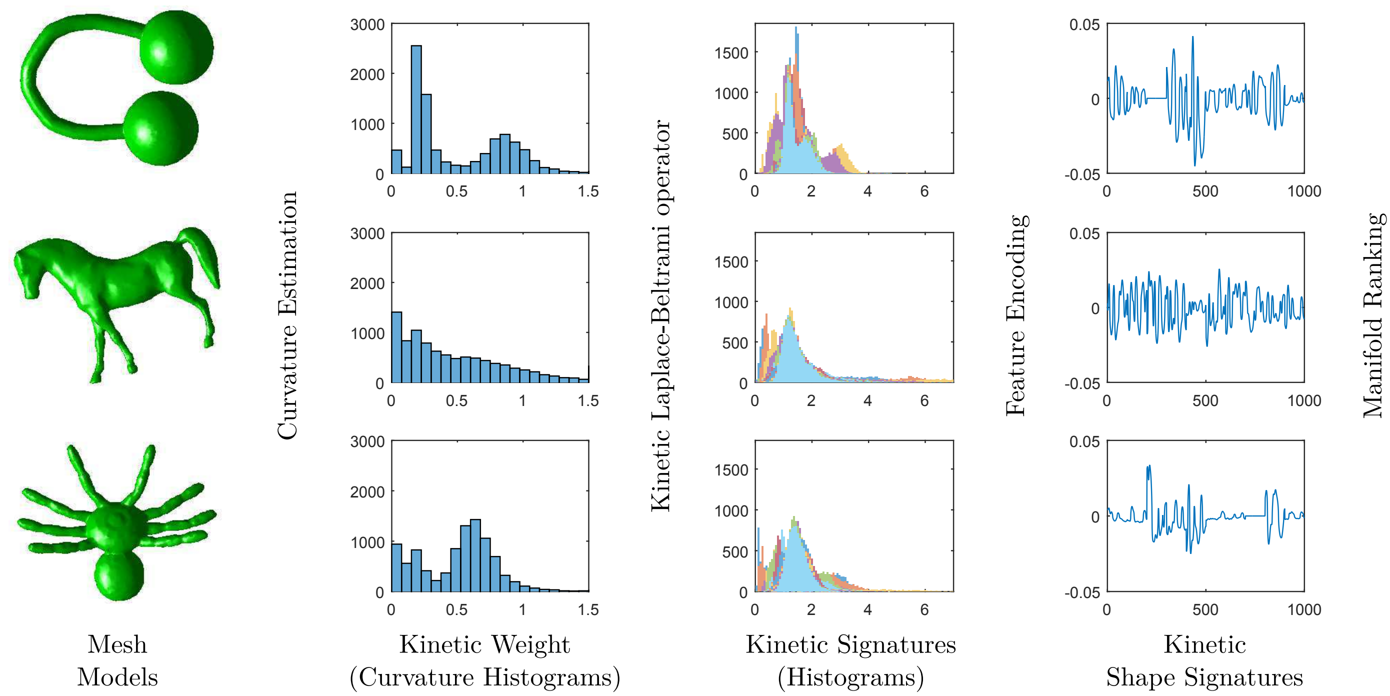


**Fig. 2. Pipeline proposed for the non-rigid shape retrieval problem for triangle meshes. By weighting the kinetic energy on the Euler-Lagrangian equation by a specific curvature term, we reduce the effect of shape articulations, causing same-class shapes' signatures to be closer to each other. Then, by encoding the kinetic signatures using either Fisher Vector or Super Vector (Section 4.1) we are able to compare shapes efficiently using Manifold Ranking technique.**

In equation (27), $\beta$ is a parameter representing the desired average curvature to rescale the curvatures and $c_x$ is the maximum principal curvature. This way, the final IWKS is given by

$$\mathrm{IWKS}(x,e) = C_e \sum_{k=1}^{\infty} \phi_k(x)^2 f_C(\Lambda_k)^2 + S(x) \tag{28}$$

Fig. 4 shows a plot of $S(\mathbf{x})$ over different deformed shapes. It is easy to see that $S(\mathbf{x})$ is less weighted on joint regions (mainly parts near leg joints) and it remains stable across deformations of the shape (see the tail and the neck).

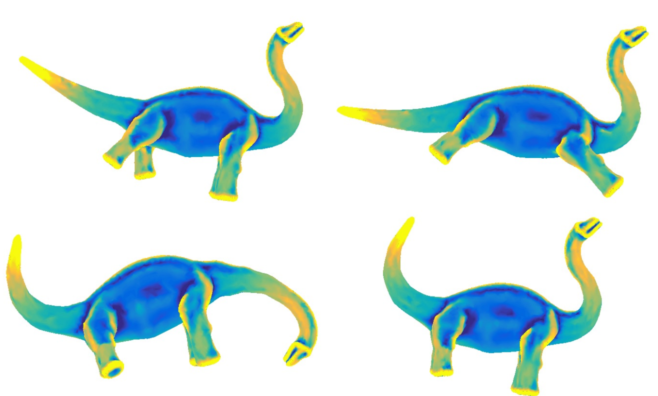

**Fig. 4. Plot of $S(\mathbf{x})$ on different dinosaur models from SHREC'15 benchmark using the same colormap. Blue stands for low values and yellow stands for high values. As can be seen, the positive curvatures remain stable along non-rigid deformations of the shape. Models are respectively 69, 171, 323 and 393.**

This modification normalizes the curvature in a more robust way, making similar shapes to have more similar curvature histograms, independent of size or sampling. For same-class shapes, histograms will have the same mean and very similar variance, while shapes from different classes will still have the same mean, however, distinct distributions. We show in Figure 5 that curvature histograms from shapes of the same class are similar even when these are articulated. The classes we selected have different curvature histograms to better illustrate the clear similarity that some classes have, however, there are many other curvature histograms which are not as similar as in the example, therefore only a histogram comparison would not be sufficient to classify those shapes. Figure 6 shows the average similarity between curvature histograms for each class in SHREC'15. It is possible to see that there is a correlation between curvature histograms of same-class shapes since there are dominant similarities in the main diagonal of Figure 6 (where we compare shapes of the same class).

In addition to using the KLBO with the IWKS, we can also use the kinetic shape spectrum to compose other spectral signatures, for example, the Heat Kernel Signature (HKS), the Scale-Invariant Heat Kernel Signature (SIHKS) and the standard Wave Kernel Signature (WKS). These computations are done exactly in the same way, changing the computation of the eigenvalues and eigenfunctions from the standard LBO to the KLBO framework. This makes all spectral signatures more robust to non-rigid deformations of the shape.

### *4.1. Encoding local spectral descriptors*

After computing local descriptors, it is necessary to encode them into shape signatures to make 3D comparisons. For this,

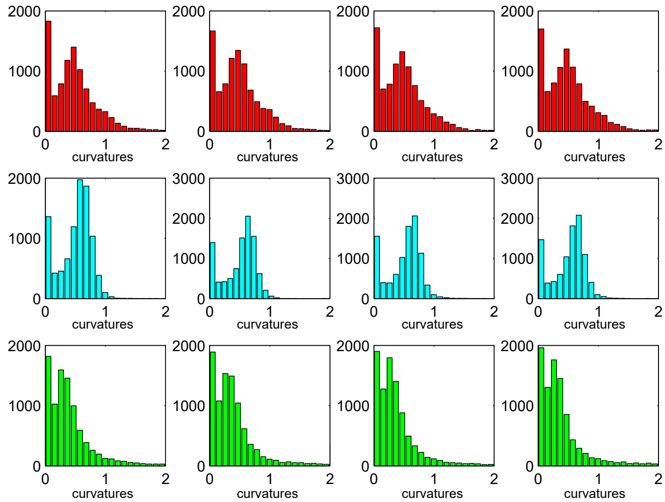


**Fig. 5. Curvature histograms of three different classes which models were articulated: *dinosaur* in the first row (red), *glasses* in the second row (cyan) and *gorilla* in the third row (green). Models are respectively 618, 323, 1007, 624, 697, 500, 1062, 312, 829, 46, 915 and 770, from SHREC'15 Lian et al. (2015). $\beta$ was set to 0.5.**

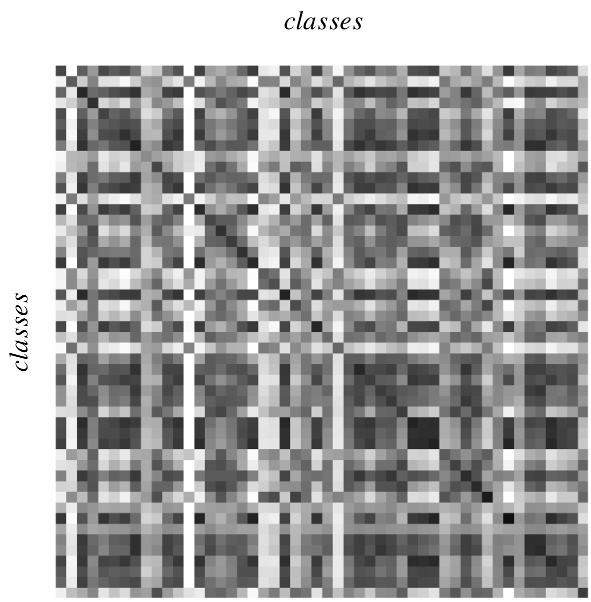


**Fig. 6. Average similarity between curvature histograms for each class in SHREC'15. Black means that the difference between classes are small, oppositely white means they are big. The distinct diagonal indicates that the within-class histograms are highly self-similar.**

we use two related methods based on the Bag-of-Words framework: the Fisher Vector from Perronnin et al. (2010) and the Super Vector from Zhou et al. (2010). Both have been applied to spectral signatures before by Limberger and Wilson (2015).

Standard Bag-of-Word methods, i.e. Histogram encoding and Kernel codebook encoding, have been first applied to shape analysis tasks by Toldo et al. (2009) and Bronstein et al. (2011). Many different schemes were proposed to improve the performance of encodings methods, e.g. Tabia et al. (2013a,b); Savelonas et al. (2016), however, we have chosen the FV and SV for their good performances in recent non-rigid shape retrieval benchmarks.

In this section, we show the mathematic formulations of the Fisher Vector and Super Vector encoding methods and then we focus in showing that the new IWKS also can be used with these encoding schemes.

FV and SV are based on the differences between descriptor means and the centers of probabilistic distribution functions (PDF) which act as a dictionary of features. These differences describe which features are and which features are not present on the object. When computing FV or SV, a Gaussian Mixture Model (GMM) is used to represent the vocabulary. Thus, the better shape features are approximated by a GMM, the more precise the encoding will be. Let $X = \{\mathbf{x}_t, \mathbf{x}_t \in \mathbb{R}^D, t = 1...T\}$ be a set of local descriptors of a shape $S$, where $T$ is the number of vertices from $S$ and $D$ the descriptor dimension, and $\lambda = \{w_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, k = 1...K\}$ a set of parameters of a GMM $p_\lambda$ (Eq. (29)), where $w_k$, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are respectively the weight, mean vector and covariance vector of the $k$-th Gaussian of a GMM.

$$p_\lambda(x) = \sum_{k=1}^{K} w_k \mathcal{N}(x|\mu_k, \Sigma_k) : \sum_{k=1}^{K} w_k = 1 \tag{29}$$

The parameters $\lambda$ can be estimated by computing the Expectation Maximization (EM) algorithm from Sanchez et al. (2013). The FV produces three-order-deviation vectors $(\mathbf{q}, \mathbf{u}, \mathbf{v})$ from the vocabulary to characterize the set of local descriptors. The first order is the association strength (soft assignment), which is computed by the posterior probability

$$q_{tk} = \frac{\exp[-\frac{1}{2}(\mathbf{x}_t - \boldsymbol{\mu}_k)^\top \Sigma_k^{-1}(\mathbf{x}_t - \boldsymbol{\mu}_k)]}{\sum_{i=1}^{K} \exp[-\frac{1}{2}(\mathbf{x}_t - \boldsymbol{\mu}_i)^\top \Sigma_i^{-1}(\mathbf{x}_t - \boldsymbol{\mu}_i)]}. \tag{30}$$

Then, the second and third orders are computed w.r.t. the mean and covariance. For each mode $k$ and each descriptor dimension $j = 1..D$, deviation vectors are computed

$$u_{jk} = \frac{1}{T\sqrt{w_k}} \sum_{i=1}^{T} q_{ik} \frac{x_{ji} - \mu_{jk}}{\sigma_{jk}}, \tag{31}$$

$$v_{jk} = \frac{1}{T\sqrt{2w_k}} \sum_{i=1}^{T} q_{ik} \left[ \left( \frac{x_{ji} - \mu_{jk}}{\sigma_{jk}} \right)^2 - 1 \right] \tag{32}$$

where $\sigma_{jk}$ are the square roots of the covariances $\Sigma_k$. In the end, FV is given by the vectorization and concatenation of the matrices $u_{jk}$ and $v_{jk}$.

$$\Gamma_{FV} = [...\mathbf{u}_k^\top ..., ...\mathbf{v}_k^\top ...]^\top \tag{33}$$

On the other hand, the SV only considers two-order-deviation vectors $(\mathbf{q}, \mathbf{u})$ but it adds a component related to the mass of each cluster $(\mathbf{s})$

$$p_k = \frac{1}{N} \sum_{t=1}^{N} q_{tk} \qquad s_k = s\sqrt{p_k}$$
$$\mathbf{u}_k = \frac{1}{\sqrt{p_k}} \sum_{t=1}^{N} q_{tk}(\mathbf{x}_t - \boldsymbol{\mu}_k) \tag{34}$$

where $s$ is a weight to balance $s_k$ and $\mathbf{u}_k$ numerically. Finally, SV is given by

$$\Gamma_{SV} = [s_1, \mathbf{u}_1^\top, ..., s_K, \mathbf{u}_K^\top]^\top \tag{35}$$

The size of the final signature depends on the parameters used to compute either the FV or SV. For FV, the final size of the descriptor ($\Gamma_{FV}$) is $2DK$ and for SV ($\Gamma_{SV}$) is $K(D+1)$.

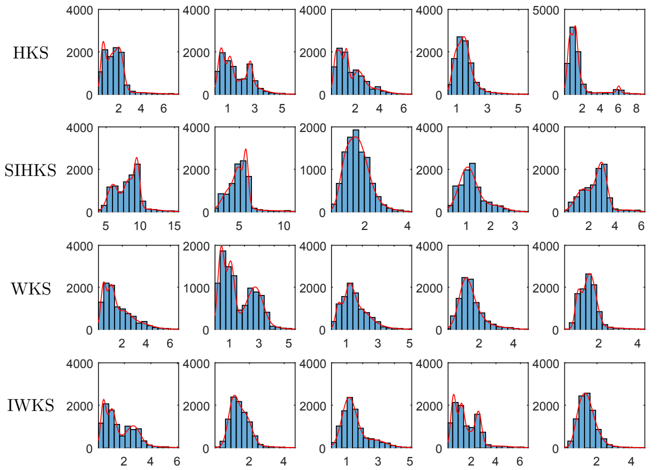


**Fig. 7. Fitting Gaussian Mixture Models to randomly-chosen shape features computed using the KLBO. Expectation-maximization algorithm was used to compute the mixture models. Each row represent features from HKS, SIHKS, WKS and IWKS.**

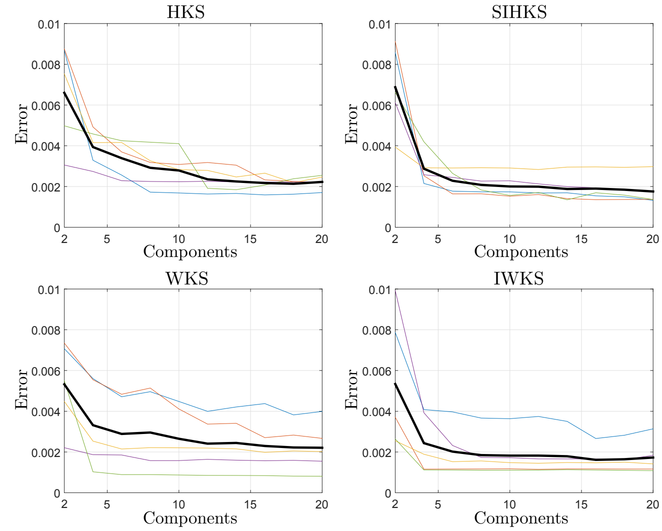


**Fig. 8. Convergence of Gaussian Mixture Models to approximate KLBO spectral descriptors. Each line of each graph represents the fitting error with different number of components on the five feature descriptors of Fig. 7. The black line represents the average loss of the five approximations. As can be seen, the error converges in most cases for the four descriptors when are used in average 5 components or more.**

In order to represent shape signatures using a GMM, it is necessary to determine whether the descriptors have the desirable characteristics (smooth histogram with a small number of peaks) to fit a GMM with a low fitting error. Therefore, we performed an empirical analysis on shape descriptors to determine the errors of fitting GMMs. We plotted histograms of descriptors frequencies (each descriptor frequency is used as a feature) and computed errors based on the differences to each bin. We show in Fig. 7, five different randomly-selected descriptor frequencies for each spectral signature (HKS, SIHKS, WKS, IWKS). In every example we fit a GMM with five components using the iterative Expectation-Maximization (EM) algorithm to show that is possible to approximate every shape feature histogram precisely, even with a small number of components.

The residuals $E$ from the GMM approximation are computed by summing all errors for each bin and normalizing by the number of observations

$$E = \sum_{i=1}^{h} \frac{|\delta_i|}{v}, \tag{36}$$

where $h$ is the number of histogram bins, $\delta_i$ is the difference from the histogram $i$-th bin value to the GMM sampled in the $x$-axis at bin's midpoint and $v$ is the number of vertices in the model.

Therefore, we plot the residuals by fitting feature histograms with different number of components. The residual plots show that the error decreases (converges to the shape feature histogram) as we increase the number of GMM components. As you can see, the approximation produces small errors which decay as we increase the number of components, thus enabling the use of GMM dictionaries with KLBO spectral signatures.

We use Efficient Manifold Ranking (EMR) algorithm from Xu et al. (2011) to compute the distances between the final FV or SV encodings. EMR works very similar to the standard Manifold Ranking algorithm. However, it is a faster version that allows out-of-sample retrieval, crucial to real-world retrieval systems. The Manifold Ranking algorithm leads to a better separation of features than using a pairwise euclidean distance by exploiting the global structure of the intrinsic manifold, created from the feature vectors. It then computes similarity between descriptors by navigating manifold graph edges, similar to a diffusion process. Therewith, a relative ranking score is assigned to each feature vector, differently from a pairwise similarity, as usually employed by dissimilarity measures. MR is becoming a standard way to compute dissimilarities in large datasets as shown in the works of Lian et al. (2010), Lian et al. (2011), Li et al. (2012), Pickup et al. (2014), Li et al. (2014b) and Furuya and Ohbuchi (2015). We used this method exactly as it is explained in Xu et al. (2011), using their implementation.

## 5. Experiments

In this section, we evaluate our descriptors in the most recent non-rigid benchmarks proposed in the literature. However, first we show that the eigenfunctions generated by the KLBO are more stable and suitable for non-rigid shape retrieval since they are less variant to non-rigid deformations.

Figure 9 show examples of eigenfunctions from three different models which were deformed by their joints. First, we show the eigenfunctions computed using the classic LBO discretization by Meyer et al. (2003). Following, we show the eigenfunctions computed using the KLBO. It is easy to see that our eigenfunctions are more stable and characterize the same regions of the shape. This makes local descriptors, first, more robust since the shape spectra is more stable and informative and, second, less variant under non-rigid transformations.

To evaluate the capability of retrieving shapes in non-rigid databases, we perform experiments on three non-rigid datasets: SHREC'10 by Lian et al. (2010), SHREC'11 by Lian et al.

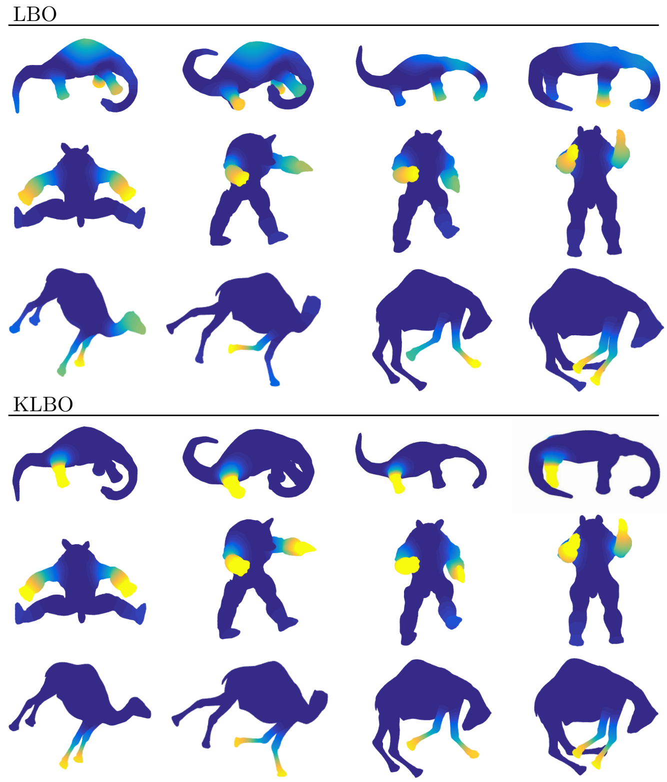


**Fig. 9. Comparison between eigenfunctions of the LBO (first three rows) and the KLBO (last three rows) of dinosaurs (6th eigenfunction), armadillos and camel (4*th* eigenfunctions) models. As you can see, the KLBO is much more stable than the LBO, and it is capable of identifying the same regions with similar weights, independently of the object's pose. This makes the KLBO more suitable for non-rigid shape retrieval.**

(2011), and SHREC'15 by Lian et al. (2015). On these benchmarks, a search query consist of using one object as query and the remaining objects as retrieval targets. This result in a dissimilarity square matrix $D$ where the entry $(i, j)$ gives the distance between models $i$ and $j$ from the database. Computing statistics over $D$ gives us knowledge of the efficiency of each method.

To compare our method against other variants, we took the state-of-the-art descriptors from each one of the datasets of Lian et al. (2010, 2011, 2015). All the results from other methods were taken from the respective author's papers. As retrieval scores, we use e-Measure (E) and mean Average Precision (mAP). The e-Measure gives a score based on the precision and recall of the first 32 retrieved models even when classes contain less than 32 models. Additionally, mAP is given by the area below the precision and recall curve, considering the plot as a square $1 \times 1$. This way, mAP varies between 0 and 1.

Table 1 shows a comparison of our best run on SHREC'10 (KLBO-SVWKS, see Table 5) against the best descriptors taken from SHREC'10 benchmark (MR-BF-DSIFT-E, DMEVD_run1, CF) and other state-of-the-art techniques, where referenced. Our descriptor exhibit the best retrieval scores when compared to all other methods. SHREC'10 dataset was one of the first datasets to deal with the problem of non-rigid shape retrieval. Some of its classes contains models that are substantially different in nature to the others. SHREC'10 is also a very challenging benchmark because some classes are very similar to each other. For these reasons, the IWKS does not performs at the top performance here since the IWKS assumes smoother transitions between models of the same class, which does not happen on this dataset for many classes.

**Table 1. Retrieval performances on SHREC'10 Non-rigid dataset.**

| Descriptor | E | mAP [%] |
|---|---|---|
| **KLBO-SVWKS** | **0.7328** | **99.1** |
| ConTopo++ (Sfikas et al. (2011)) | 0.7140 | 97.6 |
| **KLBO-SVIWKS** | 0.7137 | 93.2 |
| MR-BF-DSIFT-E | 0.7055 | 95.4 |
| DMEVD_run1 | 0.7012 | 94.1 |
| MDS-ZFDR (Li et al. (2014a)) | - | 94.1 |
| FV-IWKS (Limberger and Wilson (2015)) | 0.5867 | 82.8 |
| CF | 0.5527 | 75.2 |
| BOF-SIHKS (Bronstein et al. (2011)) | 0.5239 | 66.1 |

*Best runs from the three groups that performed better on SHREC'10 (MR-BF-DSIFT-E, DMEVD_run1, CF) and other recent descriptors that outperformed those, against our descriptor (KLBO-SVWKS). In bold are highlighted the best performances for each retrieval measure.*

**Table 2. Retrieval performances on SHREC'11 Non-rigid dataset.**

| Descriptor | E | mAP [%] |
|---|---|---|
| **KLBO-FVIWKS** | **0.7451** | **100.0** |
| 3DVFF (Furuya and Ohbuchi (2014)) | - | 99.1 |
| SD-GDM-meshSIFT | 0.7358 | 98.5 |
| MDS-ZFDR (Li et al. (2014a)) | - | 97.5 |
| SV-DSIFT (Furuya and Ohbuchi (2015)) | - | 97.2 |
| FV-IWKS (Limberger and Wilson (2015)) | 0.7318 | 97.1 |
| R-BiHDM-L23 (Ye and Yu (2015)) | 0.7300 | - |
| SGWC-BoF (Masoumi et al. (2016)) | 0.7290 | - |
| SV-LSF_kpaca50 (Furuya and Ohbuchi (2015)) | - | 96.2 |
| Geodesic Distances (LS) (Pickup et al. (2015)) | 0.7170 | - |
| MDS-CM-BOF | 0.7166 | 95.0 |
| ConTopo++ (Sfikas et al. (2011)) | 0.6950 | 94.7 |
| OrigM-n12-normA | 0.7047 | 94.4 |
| FOG+MRR | 0.6958 | 91.8 |
| BOGH | 0.6469 | 86.7 |
| LSF | 0.6327 | 85.1 |

*Best runs from the six groups that performed better on SHREC'11 (SD_GDM-meshSIFT, MDS-CM-BOF, OrigM-n12-normA, FOG+MRR, BOGH, LSF) and other recent descriptors that outperformed those, against our descriptor(KLBO-FVIWKS). In bold are highlighted the best performances for each retrieval measure.*

Table 2 compares the performance of the KLBO-FVIWKS against the best methods on SHREC'11 benchmark (SD-GDM-meshSIFT, MDS-CM-BOF, OrigM-n12-nrmA, FOG+MRR, BOGH, LSF) and other state-of-the-art techniques, where referenced. As shown, our method clearly outperforms all others, achieving an excellent retrieval score (mAP 100.0%). Our method is not based on any kind of supervised feature learning and it is able to acknowledge important characteristics of

Table 3. Retrieval performances on SHREC'15 Non-rigid dataset.

| Descriptor | E | mAP [%] |
|---|---|---|
| SV-LSF_kpaca50 (Furuya and Ohbuchi (2015)) | **0.8357** | **99.8** |
| **KLBO-FVIWKS** | 0.8269 | 99.2 |
| HAPT_run1 | 0.8150 | 97.7 |
| FV-IWKS (Limberger and Wilson (2015)) | 0.8102 | 96.9 |
| SPH_SparseCoding_1024 | 0.8047 | 96.8 |
| SGWC-BoF (Masoumi et al. (2016)) | 0.7470 | - |
| CompactBoHHKS10D | 0.7465 | 90.1 |
| SRG (Mohamed and Ben 4 (2016)) | 0.7390 | - |
| FV-WKS | 0.7242 | 87.5 |
| EDBCF_NW | 0.7076 | 85.0 |

*Best runs from the six groups that performed better on SHREC'15 (SV-LSF_kpaca50, HAPT_run1, SPH_SparseCoding_1024, CompactBoHHKS10D, FV-WKS, EDBCF_NW) and other recent descriptors in the literature against our descriptor (KLBO-FVIWKS). In bold are highlighted the best performances for each retrieval measure.*

the shapes just by inspecting its enhanced spectral components. Therefore, KLBO-FVIWKS is capable of retrieving all 19 correct matches for all 30 classes, i.e. if any object from these classes is taken as query it will retrieve all remaining shapes from the same class at first. We show in Fig. 10 an example of a very challenging class. The *snakes* example was chosen because of the difficulty that other methods that do not use diffusion geometry have to retrieve this sort of shape.

On SHREC'15 dataset, our method performs closely to the top performing method (see Table 3). Although KLBO-FVIWKS does not achieve the first position, it is very stable along other benchmarks. Differently than SV-LSF_kpaca50, which achieves a very good retrieval score on SHREC'15 but has a lower performance on SHREC'11.

We also show the performance of our method applied to another problem. We use the benchmark of Biasotti et al. (2017) to test the retrieval accuracy of similar relief patterns. Considering the entire dataset, our method is the best achieving a mAP of 0.339. In second place is LBPI with mAP 0.283. The entire dataset consist of meshes that are deformed in different ways: sampling, size, shape bending. In total, there are 15 different relief patterns (classes) and 720 different models. In Table 4 are shown different evaluation measures for the four best descriptors in this benchmark. The curvature weighting makes it possible to describe the relief patterns of the surfaces and retrieve similar patterns with a good accuracy. Retrieval statistics were taken from the benchmark paper of Biasotti et al. (2017).

Table 4. Retrieval performances on SHREC'17 Relief Patterns dataset.

| Descriptor | E | mAP [%] |
|---|---|---|
| **KLBO-FVIWKS** | **0.332** | **0.339** |
| LBPI | 0.232 | 0.283 |
| CMC-2 | 0.261 | 0.271 |
| IDAH-1 | 0.145 | 0.174 |

Table 5 shows experiments on the KLBO comparing with the classical LBO (Limberger and Wilson (2015)) to compute well known spectral signatures from the literature (HKS, SIHKS, WKS) and our IWKS. In the sixth column, we also show results when applying EMR to compute the dissimilarity matrix, differently from Euclidean distance. The last two columns show the respective improvements when applying the KLBO and EMR+KLBO over the LBO. In all cases, the method improves the results of the signatures when using EMR+KLBO, in some cases by more than 20%.

Table 6 shows detailed running times to compute KLBO. We show average times to compute one model from each database. In the second row, we show running times for the KLBO to compute the Laplacian matrix and its respective eigendecomposition. The following columns show running times for computing each spectral local descriptor. *SV* and *FV* represent the average time to compute the encoding for one model. At the last two columns it is shown the total time to compute each benchmark, using either FV or SV.

Our method is not designed for benchmarks that transform shape topologies, like deformable shapes with missing parts from Rodola et al. (2017), since we use curvatures to guide the kinetic-energy flow. Once parts of the model are removed, the curvatures are changed in the borders of the missing parts. We believe that our method could be modified to account for missing parts, using the techniques described in Rodola et al. (2017) and should perform well, but this is another research project. With a naive application of our method, we get better performance that the methods which do not explicitly account for missing parts, but not as good as those that do.

In overall, it is difficult to have one descriptor that behaves well in different kinds of data (different examples). As it can be seen, ConTopo++ performs very well in SHREC'10 however its accuracy in SHREC'11 is not in the top tier. On the other hand, our descriptors can achieve a high performance on all benchmarks tested in this paper, having more than 99% of accuracy in retrieving non-rigid shapes.

**Settings** For computing shape encodings, we use a dictionary of the first 29 models of the respective dataset. Then, we compute GMMs with 38 components. The IWKS is implemented as in Limberger and Wilson (2015) with the modification described in Eq. (28), using $m = 0.3$ and $iwksvar = 5$ for SHREC'10; $m = 0.5$ and $iwksvar = 2.5$ for SHREC'11; $m = 0.5$ and $iwksvar = 3.75$ for SHREC'15; and $m = 0.5$ and $iwksvar = 5$ for SHREC'15. We use $wksvar = 6$ in all datasets. To compute EMR, we use 100, 220, 500 and 70 landmarks, respectively, for each benchmark, using authors' implementation Xu et al. (2011). For computing shape signatures, we compute the first 300 eigenvalues for the KLBO and LBO. We evaluate local signatures (HKS,SIHKS,WKS and IWKS) with the respective parameters: HKS and SIHKS time interval $[4\ln(10)/\lambda_{300}, 4\ln(10)/\lambda_2]$ logarithmic scaled; WKS every interval $[\log(\lambda_2), \log(\lambda_{300})]$; IWKS energy interval $[\sqrt[3]{\lambda_f}, \sqrt[3]{\lambda_{300}}]$; $\lambda_i$ represent the eigenvalues and $\lambda_f$ is the first nonzero eigenvalue. Then, signatures are sampled 100 times in these intervals. In the SIHKS, we sample the first 15 frequencies after computing scale normalization. These parameters were chosen because they were the best ones when testing in a training set (a

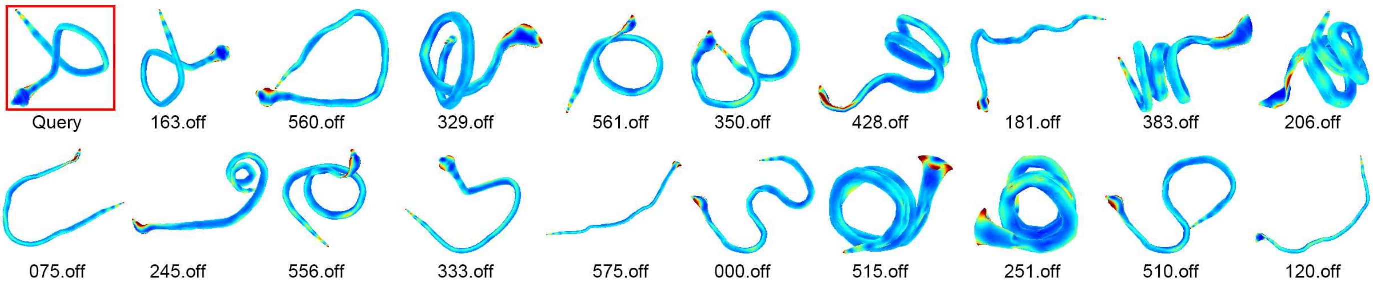


**Fig. 10. Retrieval of the model 235 (snake shown in the red square) from SHREC'11 dataset. Differently from other descriptors in Lian et al. (2010) which fail on the identification of snakes, the KLBO method is capable of identifying all 19 correct matches at first from this repository. Furthermore, the KLBO framework also retrieves all snakes given any snake as query from either SHREC'10 or SHREC'15 datasets.**

**Table 5. Experiments of the KLBO applied to different spectral descriptors and encodings on different benchmarks.**

| Benchmark | Encoding | Descriptor | LBO (e-Measure) | KLBO (e-Measure) | EMR+KLBO (e-Measure) | Improvements (KLBO) | Improvements (EMR+KLBO) |
|---|---|---|---|---|---|---|---|
| SHREC'10 | FV | HKS | 0.6570 | 0.6910 | 0.7043 | 1.0 % | 7.2 % |
| | FV | SIHKS | 0.6173 | 0.6651 | 0.7045 | 7.8 % | 14.1 % |
| | FV | WKS | **0.6738** | **0.7225** | 0.7148 | 7.2 % | 6.1 % |
| | FV | IWKS | 0.5793 | 0.6863 | 0.6767 | 18.4% | 16.8 % |
| | SV | HKS | 0.6017 | 0.6125 | 0.6878 | 1.8 % | 14.3 % |
| | SV | SIHKS | 0.6313 | 0.6776 | 0.7052 | 7.3 % | 11.7 % |
| | SV | WKS | 0.6455 | 0.7073 | **0.7328** | 9.6 % | 13.5 % |
| | SV | IWKS | 0.5957 | 0.6629 | 0.7137 | 11.3% | 19.8 % |
| SHREC'11 | FV | HKS | 0.6996 | 0.7161 | 0.7426 | 2.4 % | 6.2 % |
| | FV | SIHKS | 0.7229 | 0.7418 | 0.7425 | 2.6 % | 2.7 % |
| | FV | WKS | 0.7210 | **0.7430** | **0.7451** | 3.1 % | 3.3 % |
| | FV | IWKS | **0.7318** | 0.7420 | 0.7441 | 1.4 % | 1.7 % |
| | SV | HKS | 0.6523 | 0.6580 | 0.7361 | 0.9 % | 12.9% |
| | SV | SIHKS | 0.7189 | 0.7383 | **0.7451** | 2.7 % | 3.6 % |
| | SV | WKS | 0.7129 | 0.7425 | 0.7439 | 4.2 % | 4.4 % |
| | SV | IWKS | 0.7283 | 0.7413 | **0.7451** | 1.8 % | 2.3 % |
| SHREC'15 | FV | HKS | 0.6661 | 0.6400 | 0.7225 | -3.9 % | 8.5 % |
| | FV | SIHKS | 0.7102 | 0.7587 | 0.7988 | 6.8 % | 12.5% |
| | FV | WKS | 0.7511 | 0.7795 | 0.7925 | 3.8 % | 5.5 % |
| | FV | IWKS | **0.8102** | **0.8255** | **0.8269** | 1.9 % | 2.1 % |
| | SV | HKS | 0.5564 | 0.5514 | 0.6918 | -0.9% | 24.3% |
| | SV | SIHKS | 0.6698 | 0.7458 | 0.8019 | 11.3% | 19.7% |
| | SV | WKS | 0.6842 | 0.7452 | 0.7858 | 8.9 % | 14.9% |
| | SV | IWKS | 0.7649 | 0.8028 | 0.8232 | 5.0 % | 7.62% |

*In bold are highlighted the best retrieval performances for each benchmark and method. LBO and KLBO columns use euclidean distance to compute dissimilarities between descriptors. Improvements of KLBO over LBO are shown in the seventh column. The final improvements over the LBO descriptor due the weighting of the Kinetic energy (KLBO) and EMR are given at the last column (with respect to e-Measure).*

**Table 6. Average computation times (in seconds) for computing one signature for an average-sized model from each dataset.**

| Benchmark | KBLO | HKS | SIHKS | WKS | IWKS | FV | SV | EMR | **Total-FV** | **Total-SV** |
|---|---|---|---|---|---|---|---|---|---|---|
| SHREC'10 | 24.74 | 0.10 | 5.39 | 0.05 | 0.06 | 1.51 | 7.22 | 0.07 | **6,390** | **7,533** |
| SHREC'11 | 12.73 | 0.06 | 3.62 | 0.04 | 0.06 | 0.98 | 4.24 | 0.27 | **10,668** | **12,625** |
| SHREC'15 | 15.24 | 0.07 | 3.77 | 0.04 | 0.05 | 1.26 | 8.73 | 0.58 | **25,239** | **34,205** |

*KLBO stands for computation of curvatures, eigenvectors and eigenvalues. HKS, SIHKS, WKS and IWKS stand for time to compute respective signatures. FV and SV stand for computation times of Fisher Vector and Super Vector. EMR represents the time to perform retrieval of one model. Total times to compute signatures and retrieve all models using either FV or SV are shown in Total-FV and Total-SV columns. The average computation times of Fisher Vector approach is considerably lower because we use VLfeat implementation Vedaldi and Fulkerson (2008), while Super Vector is completely implemented on Matlab. Complexities of FV and SV are similar, thus SV would have similar computation time if it was implemented in like manner.*

subset of Lian et al. (2010)). All experiments were computed on Matlab, PC Intel Core i7 3.4GHz, 8GB RAM.

**Limitations** Joint regions, estimated by the curvature-based kinetic term, might not be precise when models exhibit high levels of noise. One solution is to blur the curvature or use different maps over the surface. For the benchmarks analyzed in this paper, noise was not a problem for computing signatures. When it comes to computation time, although our technique provides outstanding results it takes a considerable time to compute all signatures and encodings. There exist other methods which are designed to deal with time performance and scalability Sipiran et al. (2015). Another limitation of the KLBO is that most discrete LBO do not guarantee convergence when triangles in the mesh are not well-shaped Sun et al. (2009). When spectral signatures are computed from triangle meshes, they require the 3D model to have a manifold data structure, which is not easy to find in most Internet shape databases, as stated by Li Li et al. (2015). In this case, when meshes are not well-shaped, we recommend using another discretization for the Laplace-Beltrami operator, for example from Belkin et al. (2008), or reshaping the model to a watertight mesh version.

## 6. Conclusion

In this paper, we proposed to compute enhanced non-rigid spectral signatures for 3D objects. This way, we proposed the Kinetic Laplace-Beltrami operator (KLBO), based on a modification to the dynamic systems on the mesh (kinetic energy). By introducing a new curvature-based kinetic term we were able to improve significantly the retrieval performance of spectral descriptors by making energy more difficult to move in articulated regions. Furthermore, we proposed modifications to the Improved Wave Kernel signature in order to weight curvature in a more robust way, keeping it more stable over shape deformations. By combining the KLBO with spectral signatures and computing robust distances between descriptors we clearly outperform the state-of-the-art in two non-rigid benchmarks and in a relief patterns benchmark. We show that our method is consistent over different examples since it achieves excellent retrieval performances considering non-rigid databases.

## Acknowledgments

This work was supported by CAPES - Brazil, Grant No. 11892-13-7.

## References


Andreux, M., Rodola, E., Aubry, M., Cremers, D., 2014. Anisotropic laplace-beltrami operators for shape analysis, in: Sixth Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (NORDIA).

Aubry, M., Schlickewei, U., Cremers, D., 2011. The wave kernel signature: A quantum mechanical approach to shape analysis, in: Computer Vision Workshops (ICCV Workshops), IEEE International Conference on, pp. 1626–1633. doi:10.1109/ICCVW.2011.6130444.

Belkin, M., Sun, J., Wang, Y., 2008. Discrete laplace operator on meshed surfaces, in: Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry, ACM, New York, NY, USA. pp. 278–287.

Belongie, S., Malik, J., Puzicha, J., 2000. Shape context: A new descriptor for shape matching and object recognition, in: Neural Information Processing Systems, pp. 831–837.

Biasotti, S., Thompson, E.M., Aono, M., Ben Hamza, A., Bustos, B., Dong, S., Du, B., Fehri, A., Li, H., Limberger, F.A., Masoumi, M., Rezaei, M., Sipiran, I., Sun, L., Tatsuma, A., Forero, S.V., Wilson, R.C., Wu, Y., Zhang, J., Zhao, T., Fornasa, F., Giachetti, A., 2017. Shrec17 track: Retrieval of surfaces with similar relief patterns, in: 10th Eurographics workshop on 3D Object retrieval, Lyon, France.

Boscaini, D., Masci, J., Melzi, S., Bronstein, M.M., Castellani, U., Vandergheynst, P., 2015. Learning class-specific descriptors for deformable shapes using localized spectral convolutional networks, in: Proceedings of the Eurographics Symposium on Geometry Processing, Eurographics Association. pp. 13–23.

Boscaini, D., Masci, J., Rodola, E., Bronstein, M.M., Cremers, D., 2016. Anisotropic Diffusion Descriptors. Computer Graphics Forum doi:10.1111/cgf.12844.

Bronstein, A.M., Bronstein, M.M., Guibas, L.J., Ovsjanikov, M., 2011. Shape google: Geometric words and expressions for invariant shape retrieval. ACM Trans. Graph. 30, 1:1–1:20. doi:10.1145/1899404.1899405.

Bronstein, M., Kokkinos, I., 2010. Scale-invariant heat kernel signatures for non-rigid shape recognition, in: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 1704–1711. doi:10.1109/CVPR.2010.5539838.

Choukroun, Y., Shtern, A., Kimmel, R., 2016. Elliptic operator for shape analysis. CoRR arXiv:1611.01990. URL: http://arxiv.org/abs/1611.01990.

Desbrun, M., Meyer, M., Schröder, P., Barr, A.H., 1999. Implicit fairing of irregular meshes using diffusion and curvature flow, in: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA. pp. 317–324.

Elad, A., Kimmel, R., 2003. On bending invariant signatures for surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 25, 1285–1295. doi:10.1109/TPAMI.2003.1233902.

Furuya, T., Ohbuchi, R., 2014. Fusing multiple features for shape-based 3d model retrieval, in: Proceedings of the British Machine Vision Conference, BMVA Press.

Furuya, T., Ohbuchi, R., 2015. Diffusion-on-manifold aggregation of local features for shape-based 3d model retrieval. Proc. ACM International Conference on Multimedia Information Retrieval .

Grigor'yan, A., 2006. Heat kernels on weighted manifolds and applications.

Johnson, A., 1997. Spin-Images: A Representation for 3-D Surface Matching. Ph.D. thesis. Robotics Institute, Carnegie Mellon University. Pittsburgh, PA.

Levy, B., 2006. Laplace-beltrami eigenfunctions towards an algorithm that "understands" geometry, in: Proceedings of the IEEE International Conference on Shape Modeling and Applications 2006, IEEE Computer Society, Washington, DC, USA. p. 13.

Li, B., Godil, A., et al., 2012. Shrec'12 track: Generic 3d shape retrieval, in: Proceedings of the 5th Eurographics Conference on 3D Object Retrieval, Eurographics Assoc.. pp. 119–126.

Li, B., Godil, A., Johan, H., 2014a. Hybrid shape descriptor and meta similarity generation for non-rigid and partial 3d model retrieval. Multimedia Tools and Applications 72, 1531–1560.

Li, B., Lu, Y., Li, C., Godil, A., Schreck, T., et al., 2014b. Large scale comprehensive 3D shape retrieval, in: 3DOR 2014, pp. 131–140.

Li, B., Lu, Y., Li, C., Godil, A., Schreck, T., et al., 2015. A comparison of 3d shape retrieval methods based on a large-scale benchmark supporting multimodal queries. Computer Vision and Image Understanding 131, 1–27. Special section: Large Scale Data-Driven Evaluation in Computer Vision.

Li, C., BenHamza, A., 2013. A multiresolution descriptor for deformable 3d shape retrieval. The Visual Computer 29, 513–524. doi:10.1007/s00371-013-0815-3.

Li, P., Ma, H., Ming, A., 2016. A non-rigid 3d model retrieval method based on scale-invariant heat kernel signature features. Multimedia Tools and Applications , 1–24doi:10.1007/s11042-016-3606-9.

Lian, Z., Godil, A., et al., 2010. SHREC'10 Track: Non-rigid 3D Shape Retrieval, in: Daoudi, M., Schreck, T. (Eds.), Eurographics Workshop on 3D Object Retrieval, The Eurographics Assoc.

Lian, Z., Godil, A., et al., 2011. SHREC'11 track: shape retrieval on non-rigid 3D watertight meshes, in: 3DOR 2011, Eurographics Assoc.. pp. 79–88.

Lian, Z., Godil, A., Sun, X., Xiao, J., 2013. CM-BOF: visual similarity-based

3d shape retrieval using clock matching and bag-of-features. Machine Vision and Applications 24, 1685–1704.

Lian, Z., Zhang, J., et al., 2015. Non-rigid 3D Shape Retrieval, in: 3DOR 2015, The Eurographics Assoc.

Limberger, F.A., Wilson, R.C., 2015. Feature encoding of spectral signatures for 3D non-rigid shape retrieval, in: Proceedings of the British Machine Vision Conference (BMVC), BMVA Press. pp. 56.1–56.13.

Litman, R., Bronstein, A., Bronstein, M., Castellani, U., 2014. Supervised learning of bag-of-features shape descriptors using sparse coding. Symposium on Geometry Processing (SGP) .

Litman, R., Bronstein, A.M., 2014. Learning spectral descriptors for deformable shape correspondence. IEEE Transactions on Pattern Analysis and Machine Intelligence 36, 171–180.

Masoumi, M., Li, C., Ben Hamza, A., 2016. A spectral graph wavelet approach for nonrigid 3d shape retrieval. Pattern Recognition Letters .

Melzi, S., Rodolà, E., Castellani, U., Bronstein, M.M., 2017. Localized manifold harmonics for spectral shape analysis. CoRR arXiv:1707.02596. URL: http://arxiv.org/abs/1707.02596.

Meyer, M., Desbrun, M., Schroder, P., Barr, A., 2003. Discrete differential-geometry operators for triangulated 2-manifolds, in: Hege, H.C., Polthier, K. (Eds.), Visualization and Mathematics III. Mathematics and Visualization, pp. 35–57. doi:10.1007/978-3-662-05105-4_2.

Mohamed, W., Ben 4, A., 2016. Deformable 3d shape retrieval using a spectral geometric descriptor. Applied Intelligence , 1–17doi:10.1007/s10489-015-0746-y.

Perronnin, F., Sánchez, J., Mensink, T., 2010. Improving the fisher kernel for large-scale image classification, in: Proceedings of the 11th European Conference on Computer Vision: Part IV, Springer-Verlag, Berlin, Heidelberg. pp. 143–156. URL: http://dl.acm.org/citation.cfm?id=1888089.1888101.

Pickup, D., Sun, X., Rosin, P.L., Martin, R.R., 2015. Euclidean-distance-based canonical forms for non-rigid 3d shape retrieval. Pattern Recognition 48, 2500 – 2512.

Pickup, D., Sun, X., Rosin, P.L., Martin, R.R., Cheng, Z., Lian, Z., et al., 2014. SHREC'14 track: Shape Retrieval of Non-Rigid 3D Human Models, in: Proceedings of the 7th Eurographics workshop on 3D Object Retrieval, Eurographics Assoc.

Pickup, D. and, S.X., Rosin, P.L., Martin, R.R., 2016. Skeleton-based canonical forms for non-rigid 3d shape retrieval. Computational Visual Media , 1–13doi:10.1007/s41095-016-0045-5.

Pottmann, H., Wallner, J., Huang, Q.X., Yang, Y.L., 2009. Integral invariants for robust geometry processing. Comput. Aided Geom. Des. 26, 37–60. URL: http://dx.doi.org/10.1016/j.cagd.2008.01.002, doi:10.1016/j.cagd.2008.01.002.

Reuter, M., Wolter, F.E., Peinecke, N., 2005. Laplace-spectra as fingerprints for shape matching, in: Proceedings of the 2005 ACM Symposium on Solid and Physical Modeling, ACM, New York, NY, USA. pp. 101–106. doi:10.1145/1060244.1060256.

Reuter, M., Wolter, F.E., Peinecke, N., 2006. Laplace-beltrami spectra as 'shape-dna' of surfaces and solids. Comput. Aided Des. 38, 342–366. doi:10.1016/j.cad.2005.10.011.

Rodola, E., Cosmo, L., Litany, O., Bronstein, M.M., Bronstein, A.M., Audebert, N., Ben Hamza, A., Boulch, A., Castellani, U., Do, M.N., Duong, A.D., Furuya, T., Gasparetto, A., Hong, Y., Kim, J., Saux, B.L., Litman, R., Masoumi, M., Minello, G., Nguyen, H.D., Nguyen, V.T., Ohbuchi, R., Pham, V.K., Phan, T.V., Rezaei, M., Torsello, A., Tran, M.T., Tran, Q.T., Truong, B., Wan, L., Zou, C., 2017. Deformable Shape Retrieval with Missing Parts, in: Pratikakis, I., Dupont, F., Ovsjanikov, M. (Eds.), Eurographics Workshop on 3D Object Retrieval, The Eurographics Association. doi:10.2312/3dor.20171057.

Rusinkiewicz, S., 2004. Estimating curvatures and their derivatives on triangle meshes, in: Proceedings of the 3D Data Processing, Visualization, and Transmission, 2Nd International Symposium, IEEE Computer Society, Washington, DC, USA. pp. 486–493.

Rustamov, R.M., 2007. Laplace-beltrami eigenfunctions for deformation invariant shape representation, in: Proceedings of the Fifth Eurographics Symposium on Geometry Processing, Eurographics Assoc.. pp. 225–233.

Sanchez, J., Perronnin, F., Mensink, T., Verbeek, J., 2013. Image classification with the fisher vector: Theory and practice. International Journal of Computer Vision 105, 222–245.

Savelonas, M.A., Pratikakis, I., Sfikas, K., 2016. Fisher encoding of differential fast point feature histograms for partial 3d object retrieval. Pattern Recogn. 55, 114–124. URL: https://doi.org/10.1016/j.patcog.2016.02.003, doi:10.1016/j.patcog.2016.02.003.

Sfikas, K., Pratikakis, I., Theoharis, T., 2011. Contopo: Non-rigid 3d object retrieval using topological information guided by conformal factors, in: Proceedings of the 4th Eurographics Conference on 3D Object Retrieval, Eurographics Assoc.. pp. 25–32.

Sipiran, I., Bustos, B., Schreck, T., Bronstein, A.M., Bronstein, M., Castellani, U., Choi, S., Lai, L., Li, H., Litman, R., Sun, L., 2015. Scalability of Non-Rigid 3D Shape Retrieval, in: 3DOR 2015, The Eurographics Assoc.

Sun, J., Ovsjanikov, M., Guibas, L., 2009. A concise and provably informative multi-scale signature based on heat diffusion, in: Proceedings of the Symposium on Geometry Processing, Eurographics Association. pp. 1383–1392.

Tabia, H., Picard, D., Laga, H., Gosselin, P.H., 2013a. 3D Shape Similarity Using Vectors of Locally Aggregated Tensors, in: IEEE International Conference on Image Processing, Melbourne, Australia. pp. 2694–2698. URL: https://hal.archives-ouvertes.fr/hal-00832182. iEEE International Conference on Image Processing 2013.

Tabia, H., Picard, D., Laga, H., Gosselin, P.H., 2013b. Compact Vectors of Locally Aggregated Tensors for 3D shape retrieval, in: Eurographics Workshop on 3D Object Retrieval, Girona, Spain. pp. –. URL: https://hal.archives-ouvertes.fr/hal-00807501.

Taubin, G., 1995. A signal processing approach to fair surface design, in: Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques, ACM, New York, NY, USA. pp. 351–358.

Toldo, R., Castellani, U., Fusiello, A., 2009. Visual vocabulary signature for 3d object retrieval and partial matching, in: Proceedings of the 2Nd Eurographics Conference on 3D Object Retrieval, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland. pp. 21–28. URL: http://dx.doi.org/10.2312/3DOR/3DOR09/021-028, doi:10.2312/3DOR/3DOR09/021-028.

Vedaldi, A., Fulkerson, B., 2008. VLFeat: An open and portable library of comp. vision algorithms. http://www.vlfeat.org/.

Xu, B., Bu, J., Chen, C., Cai, D., He, X., Liu, W., Luo, J., 2011. Efficient manifold ranking for image retrieval, in: Proceedings of the 34th Annual International ACM SIGIR Conference (SIGIR'11).

Xu, G., 2004. Discrete laplace–beltrami operators and their convergence. Comput. Aided Geom. Des. 21, 767–784.

Xu, G., 2006. Discrete laplace-beltrami operator on sphere and optimal spherical triangulations. Int. J. Comput. Geometry Appl 16, 75–93.

Ye, J., Yu, Y., 2015. A fast modal space transform for robust nonrigid shape retrieval. The Visual Computer , 1–16.

Zhou, X., Yu, K., Zhang, T., Huang, T.S., 2010. Image classification using super-vector coding of local image descriptors, in: Proceedings of the 11th European Conference on Computer Vision: Part V, Springer-Verlag, Berlin. pp. 141–154.